AUTOMATIC POINT MATCHING BETWEEN SAR AND OPTICAL IMAGES, MATCH-POINT QUALITY CHARACTERISATION AND ERROR DETECTION

Soukal Peter, Baltsavias Emmanuel

Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland
(psoukal@student.ethz.ch , manos@geod.baug.ethz.ch)

1. INTRODUCTION, AIMS AND INPUT DATA

The primary aim of this work is to do automatic point transfer via image matching between SAR and optical images. In our concrete application, the SAR images are orthorectified and have an underlying DSM. SAR orthorectification can be performed, with few exceptions, especially for older sensors, without GCPs. The point transfer then provides GCPs for optical image georeferencing, avoiding the costly and difficult GCP acquisition by other means. About 30-50 (usually well-distributed) points are needed, half of them as GCPs, the rest as check points. The same procedure is needed for other purposes, e.g. coregistration of SAR and optical images (e.g. in fast response applications, as post-hazard mapping), where optical and SAR images need to be fused, whereby optical images provide better interpretability and usually resolution, while SAR weather and day/night independence. SAR and optical images are spectrally and geometrically very different and have multitemporal and other (e.g. noise characteristics) differences, making matching extremely challenging. Only a very small % of the matches (0.01 to 1%) is correct. Thus, we need to derive quality values for each match-point and based on that do error detection and keep only the N best points needed. This finally turned out to be the most important and challenging part of the work, and not the matching itself. It also has a much larger significance, as the same quality characterization and error detection methods can be used in matching of optical images for automatic DSM generation, or matching of SAR images for other purposes, using the same matching method as here, and with adaptations also for other matching methods. This can solve the problem of deriving quality values for each match-point, which is essential not only for error detection and reliable automated processing, but also other applications like data (e.g. DSM) fusion. Previous research work performed on the paper topic is very limited (see [1], [2], [3], [5]), while we do not know of any commercial software providing this functionality, at least reliably.
The 15 x 15 km^2 test area, close to the city of Thun, CH, had an elevation range of 1700m, many landcover classes and was imaged by 5 SAR and optical datasets each. These were combined in pairs, such that the GSD of SAR-Optical images is similar. The major image parameters are listed in Table 1.

Table 1. Major image parameters. For each platform/sensor in brackets, the band and GSD.

SAR Images	SAR Acquisition Date	Optical Images	Optical Acquisition Date
TerraSAR (X, 3m)	10.2008	SPOT 5/HRG (PAN, 2.5m)	8.2005
TerraSAR (X, 3m)	10.2008	ALOS/PRISM (PAN, 2.5m)	1.2006
RADARSAT-1 (C, 7m)	2.2008	LANDSAT 7 (PAN, 15m)	3.2000
PALSAR (L, 8m)	11.2008	LANDSAT 7 (PAN, 15m)	3.2000
ERS-1 (C, 25m)	4.2008	IRS-P6 (NIR, 23.5m)	7.2006
ASAR (C, 25m)	4.2008	IRS (NIR, 23.5m)	7.2006
ERS-1 (C, 25m)	4.2008	LANDSAT 7 (R, 30m)	6.2001
ASAR (25m)	4.2008	LANDSAT 7 (R, 30m)	6.2001

Results mentioned below refer only to the RADARSAT-1 and LANDSAT 7-PAN pair, due to lack of space.

2. METHODOLOGY

The above SAR-Optical pairs were processed using the steps: Preprocessing to reduce noise (for SAR espec. speckle) ; Coarse co-registration of optical to SAR images (performed manually with a polynomial transformation and remaining max parallaxes in this case of ca. 32 pels) ; Generation of image pyramids (using a 3x3 Gaussian filter, 5 pyramid levels and resolution reduction factor 2, so that max parallaxes on top level are 2 pels, the latter required as approximations for the specific matching method used) ; At each (pyramid) level, Wallis filtering for optical images to enhance, espec. in shadow areas, contrast and get more edges (not for SAR, as remaining speckle noise is large and non-salient edges would be much enhanced), Computation of edge gradients (to increase similarity between SAR and optical images compared to original grey values) and Adaptive thresholding of edge gradients (to reduce non-salient edges due to remaining noise) ; Elimination of match-points with no or very weak texture (as with poor x- or y-gradients, matching will not work) ; At each level, Least Squares Matching (LSM) and computation of quality measures, Match-point quality characterisation, Error detection and elimination, and Derivation of approximations for the lower level ; Evaluation of the results at level 0 (original images). Some details (limited due to lack of space) on some of the above steps are listed below.

Noise reduction includes an adaptive local method for optical images and anisotropic diffusion for SAR ones. Elimination of poor texture points before matching leads to 2-13% less match-points, increasing with lower level.

LSM (see [4]) was used, as it is very accurate and provides many quality criteria, even if slow and requiring good approximations (a few pixels for the parallaxes, i.e. shifts of the affine transformation used in LSM). Patches of 13x13 pels from the SAR image were used as reference and the optical image patches (search) were affine transformed to them in an iterative fashion, minimizing the squared sum of residuals (here edge gradient differences). All image points, apart a small image border, were matched. Although LSM is an area-based matching (ABM) method, using gradients for matching, makes it a hybrid approach, i.e. something between ABM and feature-based matching (FBM), where here features are edges. Preliminary investigations extracting points or edges with interest operators or edge detectors, showed that a classical FBM would be more difficult than our

approach, as points are too many, not distinct and very different between SAR and optical images, while one-pixel-wide edges are not only different and less than our edge gradients but also introduce false parallaxes. The 11 quality measures provided by LSM are: no. of iterations, normalized cross-correlation computed separately after iteration convergence, sigma a posteriori, estimated values of 6 affine parameters and the standard deviations of the shifts (divided by sigma a posteriori). They are checked whether they follow a normal distribution, needed for deriving quality measure thresholds, and this was indeed the case, with the exception of the no. of iterations, and much less some skewness in the normal distribution, caused by much more poor than good match-points in this matching application. Regarding these 11 quality measures, poor match-points are on one side of the distribution, except for the 6 affine parameters, which are two-sided, and the no. of iterations, which for one iteration are also rather suspicious. Regarding error detection after matching, we optionally use at each level the following tests:

1. Bi-directional matching consistency (matching search to reference patch and opposite, with a difference of more than 1 pel between the two matchings for the reference point leading to rejection), points rejected 39 - 72%, increasing as levels decrease.
2. Combination of quality measures. Each measure should "pass" in an OR fashion, whereby thresholds for each measure are derived using robust statistics, median and the robust approximation of the standard deviation, s(MAD), with MAD = median absolute deviation from the median and s(MAD) = 1.484 * MAD, and a threshold for each measure derived as (median +/- n * s(MAD)), with n derived based on long a priori knowledge and empirical tests. Points rejected after this test are 99.3- 93.4%, decreasing as levels decrease.
3. The measures of test 2 above are each given a normalised score (0-1) using a separate fuzzy function. Then, a single quality value is computed using a weighted averaging of each quality measure score, whereby weights for each criterion are determined empirically. This allows selecting the N best match-points, in this work the best 100.
4. At the end of the tests, the local consistency of the x- and y-parallax of a point is tested and if enough neighbours exist, the point is deleted or its parallaxes are corrected to the median parallaxes of the neighbourhood.

We evaluated two test combinations (1 and 2). Combination 1: tests 1, 2, 4 and Combination 2: test 3 (we tried also test 1 after 3 but too many points, many correct, were rejected, indicating the ad-hoc nature of this test).

The above tests can be performed in the whole image or in tiles to ensure good point distribution.

Due to lack of space, we do not treat details of derivation of approximations for parallaxes (shifts of affine transformation), which are very crucial, especially for LSM, and influence how many pyramid levels are needed, how matching results are propagated if more than one level needed, and the number of match points at 0 level. As a short statement, we use for approximations' propagation only the GOOD match-points (passing the error detection tests at each level), with a nearest neighbour interpolation of the parallaxes for the lower level points.

For the second top level, we derive approximations and match ALL points, for lower levels only points within a neighbourhood (here 9x9 pels) of each good match-point of the upper level.

3. RESULTS

Of the above 2 test combination methods, the 2nd is much more preferable and discussed below. It is faster (does not need the time-consuming bi-directional matching), provides more reliable results, a single quality value per point and gives control over the N best points kept. To ensure a good point distribution, for combination method 2, we select the best 100 points on level 0 plus the best point in image tiles (here 3x3) that are empty. The evaluation of the results is very difficult as there are no reference data (can hardly be on image level, contrary to reference DSMs) and manual measurements are mostly even worse than matching results due the large differences between the matched images. Thus, we use a visual inspection of match-points and a classification in % of correct, unsure and false matchings, which for the RADARSAT-1 / LANDSAT 7-PAN were 98%, 2% and 0% respectively. This is a very good result but based on additional PRELIMINARY investigations, we state that: PALSAR / LANSAT 7-PAN results were slightly even better (almost perfect), C-band sensors other than RADARSAT-1, like ERS-1 and ASAR, give similar results among them, but clearly worse than using RADARSAT-1, while TerraSAR has not been processed yet. A visual inspection of combination 1 showed only 20% correct points for the RADARSAT-1 / LANDSAT 7-PAN pair.

4. DISCUSSION AND CONCLUSIONS

The above results show that our approach delivers very good results, although the test data was very variable and difficult. The quality characterization and error detection for matching are innovative. Further research may include refinement of the above methods for error detection (use DSM slope and roughness, landcover and a subsequent (robust) sensor orientation), estimation of absolute single quality values (expressed here in pixels), avoidance of manual coarse coregistration (using approximate sensor position and attitude data, espec. RPCs) and tests of these methods in matching of only optical or SAR images and for other matching methods and associated quality criteria.

5. REFERENCES

[1] Christophe, E., Inglada, J., Giros A., 2008. Orfeo toolbox: a complete solution for mapping from high resolution satellite images. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 37, Part B4, pp. 1263-1268.

[2] Dare, P., Dowman, I., 2001. An improved model for automatic feature-based registration of SAR and SPOT images. ISPRS Journal of Photogrammetry and Remote Sensing 56 (1): 13-28.

[3] Goncalves, J., Dowman, I., 2002. Precise orientation of SPOT panchromatic images with tie points to a SAR image. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 34, Part 3A, pp. 125-130.

[4] Gruen A., Baltsavias, E.P., 1988. Geometrically Constrained Multiphoto Matching. Photogrammetric Engineering & Remote Sensing (54(5): 633 - 641.

[5] Reinartz, P., Müller, R., Schwind, P., Suri, S., Bamler, R., 2011. Orthorectification of VHR optical satellite data exploiting the geometric accuracy of TerraSAR-X data. ISPRS Journal of Photogrammetry and Remote Sensing 66 (1): 124-132.